



PROPOSERS DAY

December 8, 2021

Briefings begin at 10:00 EST





ASKEM - Agenda

Proposers Day – December 8, 2021

TIME (ET)	EVENT	SPEAKER
9:00 AM - 10:00 AM	Online Registration	
10:00 AM - 10:05 AM	Security Briefing	Mr. Troy Blackburn
10:05 AM - 10:20 AM	Human Use Briefing	Ms. Lisa Mattocks <i>DARPA HSR</i>
10:20 AM - 10:45 AM	Contracts Management Office Briefing	Ms. Jennifer Mack
10:45 AM - 11:30 AM	Automating Scientific Knowledge Extraction and Modeling (ASKEM) Presentation	Dr. Joshua Elliott <i>Program Manager, DARPA I2O</i>
11:30 AM - 11:45 AM	Submit Questions	
11:45 AM - 1:15 PM	PM Question Review and Informal Teaming Discussions	
1:15 PM - 2:00 PM	Government responses to questions	(Answer attendee questions)



Welcome!

Automating Scientific Knowledge Extraction and Modeling (ASKEM)

Proposers Day

IS

UNCLASSIFIED

Public Releasable Info Only

I2O Program Manager: Joshua Elliott

December 8th 2021





Rules of Engagement for virtual meetings

Level of meeting:

- Meeting is **Unclassified**
- NO CUI

Prohibited:

- **Video and/or audio recording of this meeting is prohibited**



Denice Holden
Program Security Officer (PSO)
(571) 218-4277

Troy Blackburn
Program Security Representatives (PSR)
(571) 218-4626

Human Subjects Research at DARPA

Lisa Mattocks
I2O ADPM/HSR Action Officer





Definition of Human Subjects Research in Federal and DoD Policies

Use of Human Subjects in an activity constituting a systematic investigation designed to develop and/or contribute to generalizable knowledge is considered Human Subjects Research, where:

The term "*human subject*" means a living individual about whom an investigator (whether professional or student) conducting research:

- (i) Obtains information or biospecimens through intervention or interaction with the individual, and uses, studies, or analyzes the information or biospecimens; or
- (ii) Obtains, uses, studies, analyzes, or generates identifiable private information or identifiable biospecimens.

Intervention: includes both physical procedures by which information or biospecimens are gathered (*e.g.*, venipuncture) and manipulations of the subject or the subject's environment that are performed for research purposes.

Interaction: includes communication or interpersonal contact between investigator and subject.

Private information: includes information about behavior that occurs in a context in which an individual can reasonably expect that no observation or recording is taking place, and information that has been provided for specific purposes by an individual and that the individual can reasonably expect will not be made public (*e.g.*, a medical record).

Identifiable private: information is private information for which the identity of the subject is or may readily be ascertained by the investigator or associated with the information.

Identifiable biospecimens: is a biospecimen for which the identity of the subject is or may readily be ascertained by the investigator or associated with the biospecimen.

Any DARPA-funded research which involves humans as defined on this page MUST be considered HSR.



All DARPA human subjects research protocols must go through **two** reviews

1st review

Local Level (local IRB)



2nd review

DoD HRPO Level (Administrative Review)



HSR approval Process

- Principal Investigator submits protocol to local IRB for review and approval
- HSR package is then submitted for DoD HRPO review and approval
 - Includes local IRB approval letter
 - Federal Wide Assurance (of institution performing research)
 - Informed Consent Document ***Make sure informed consent document includes statement that the research is being funded by DoD and thus the DoD has access to the data***
 - Recruiting Materials
 - Biosketches/CVs
 - Training Certifications
- DoD HRPO reviews entire package
 - May go back to PI with comments/recommendations/changes
- Once DoD HRPO gives approval, HSR research can begin
- Note that protected populations (i.e. military, pregnant women, etc) have special regulations that need to be followed. This includes such things as command level approval for recruitment of subjects (active duty military)

Note – DoD HRPO review and approval can take anywhere from 3-6 months. Do not delay in starting this process!



- **If possible, submit an IRB approval letter and/or a Draft HSR Protocol with proposal.** Especially, in cases where humans are involved and you don't know that the work is really HSR. Having an IRB already look at it will help you and DARPA in moving forward faster.
- **If you do not have an internal IRB, you have one of three options**
 - Hire a commercial IRB
 - Work with the Contracting Agent to determine if they have an internal IRB that could assist
 - If work involves collaboration with other performers, considering using their IRB
- **If you have a contract involving subcontractors who are conducting HSR; they will also need to obtain HSR approval.** Any performer including subcontractors must receive HSR approval through the local IRB and the DoD HRPO before start of their research.
- **If you make changes to the statement of work, they also need to be approved.** If the changes are to the HSR portion of the work, the revisions will have to go through the local IRB for review, as well as DoD HRPO review for approval and concurrence.



Points of Contact

Government HSR and ASR Action Officer

Ms. Lisa Mattocks
Lisa.Mattocks@darpa.mil
571-218-4424

HSR POC

Dr. Sahar Zafar
Sahar.Zafar.ctr@darpa.mil
703-362-3438

Automating Scientific Knowledge Extraction and Modeling (ASKEM)

Joshua Elliott
Program Manager, DARPA/I2O

Proposers Day

December 8, 2021





ASKEM Objective

Develop tools for agile creation and sustainment of complex models and simulators that support decision making in rapidly evolving mission and scientific domains such as viral epidemics like COVID-19 and the impacts of space weather

AI-assisted modeling tool-chain that is responsive to the pace and scale of modern data-intensive decision cycles



Problem: Simulators, models, and knowledge synthesis fail

An example from COVID-19

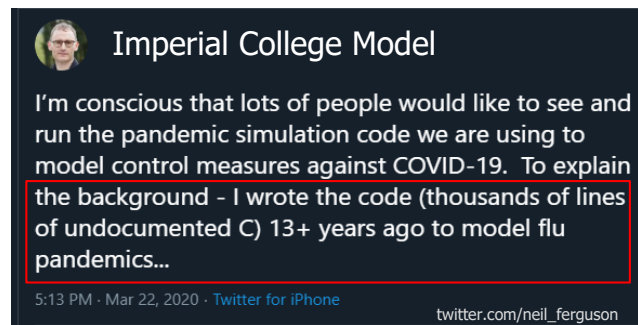
Simulator tools



- Dozens of black box simulators
- No traceability to model to provide explanations

“For accuracy of prediction, all models fared very poorly.” – *Chin et al.*¹ – August 2020

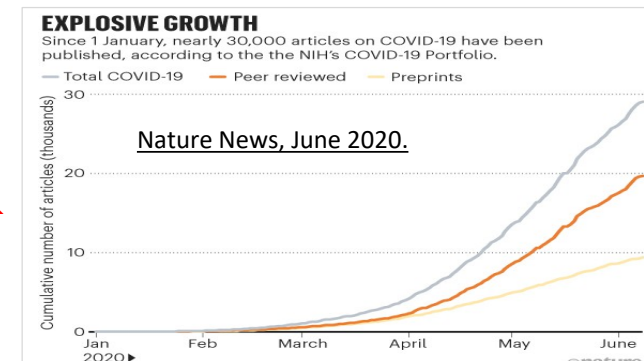
Modeling tools



- Semantically opaque legacy models
- No traceability to source knowledge

“The model...is a ‘buggy mess that looks more like a bowl of angel hair pasta than a finely tuned piece of programming’” – *David Richards*² – May 2020

Knowledge synthesis tools



- Pace and scale of new research outstripped capability to update models

“Never before has there been an explosion of research literature like that taking place this year on COVID-19.” – *Porter et al.*³ – Nov 2020

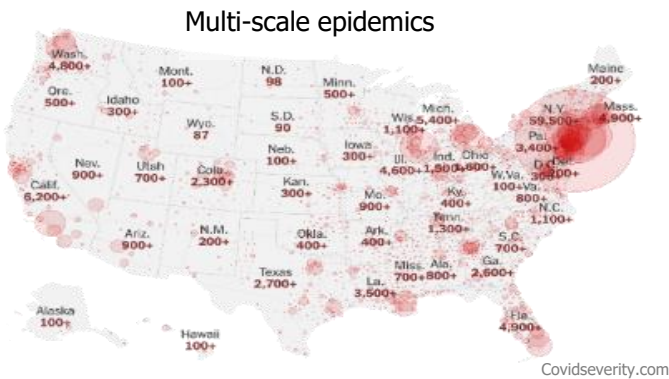
These same issues (and more) pervade all domains of scientific modeling

1) A case study in model failure? COVID-19 daily deaths and ICU bed utilisation predictions in New York state. *Eur J Epidemiol* **35**, 733–742 (2020).

2) Telegraph.co.uk/technology/2020/05/16/coding-led-lockdown-totally-unreliable-buggy-mess-say-experts/; 3) Porter et al. Tracking and Mining the COVID-19 Research Literature. *Front. Res. Metr. Anal.*, Nov 2020.

Epidemiology and pandemic response

- Army C5ISR
- NIH/CDC/NSF/DOD
- AFRICOM J26
- Gates Foundation



100s open-source models:
github.com/ihmeuw; Other sources from
 academic and gov partners

Rationale for ASKEM

- **Pace and scale** of published knowledge requires iteration
- **Legacy code** difficult to understand, adapt, debug
- **Black box** simulators have no explainability
- Types: ODE/PDE*, ABM**, or Optimization

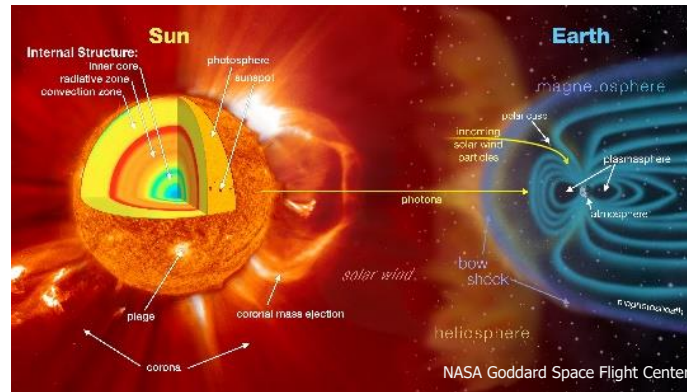
Benefits to partners: policy analysis in response to rapidly
 changing and unexpected events

*ODE/PDE: Ordinary and Partial Differential Equations

**ABM: Agent-based Models

Space Weather

- NASA/NOAA/USGS
- EPRI
- USSF/ASAF
- NGA



10s of open-source models:
 UM Space Weather Modeling Framework:
<https://github.com/MSTEM-QUA>

Rationale for ASKEM

- Requires rapid iteration for **timely decision support**
- Adds high-res **heterogeneity** in space *and* time
- Types: ODE/PDE and empirical ML

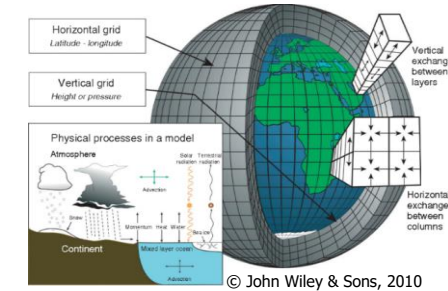
Benefits to partners: rapid recalibration of models when
 data sources change, improve multi-model coupling

Potential Communities of Practice

NIH: Viral epidemics (NIAID),
 Cancer (NCI), BRAIN initiative

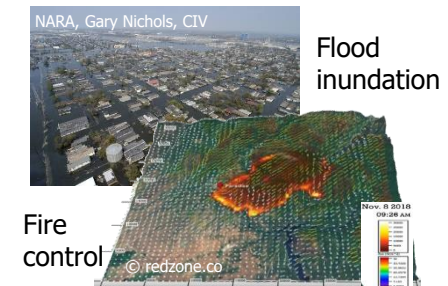
Climate Modeling

- OUSD-P Strategy and Force Development
- NIC Climate Security Advisory Council
- DOE Advanced Scientific Computing Research
- NOAA Research



Fire response and flood prevention

- USACE Flood Risk Management Program
- DHS National Risk Management Center
- USDA National Interagency Fire Center and Fire Fuel and Smoke (FFS) Science Program



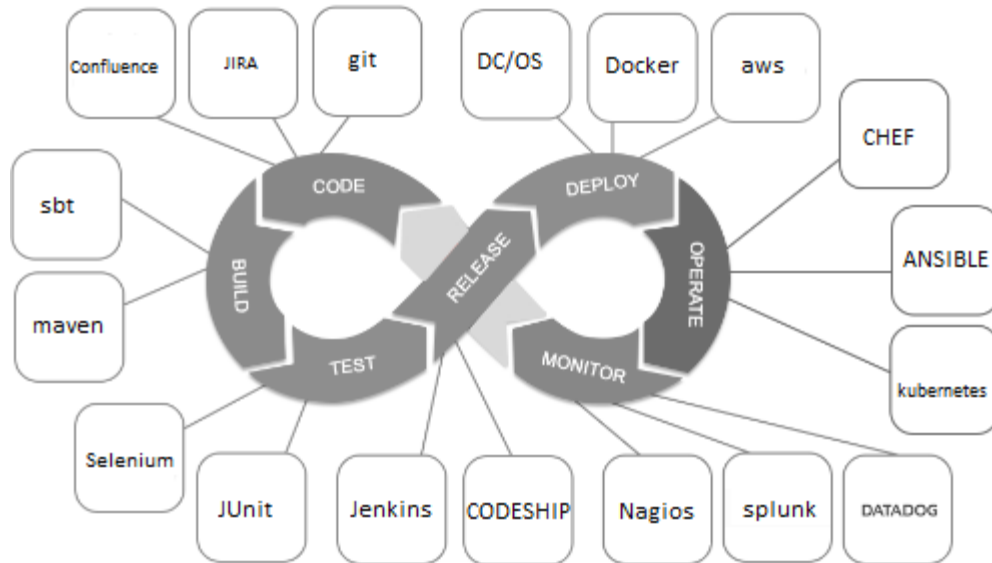


Vision: A DevOps-style approach to scientific modeling and knowledge sharing

The creation, validation, distribution, and application of scientific knowledge is a continuous iterative process, similar to software development

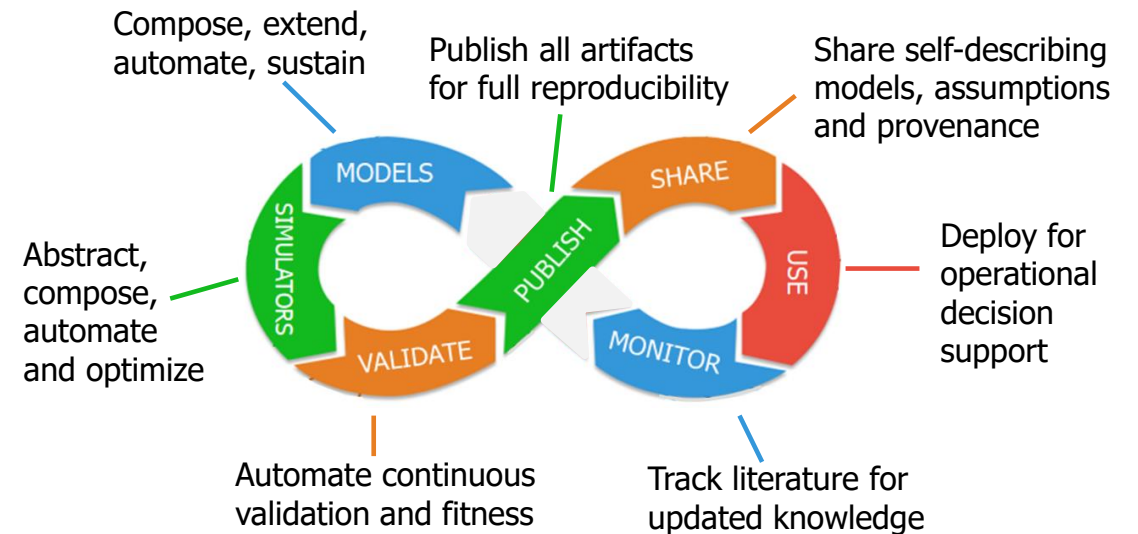
DevOps for software

Enables continuous, traceable and verifiable improvement, empowered by a **new class of tools**



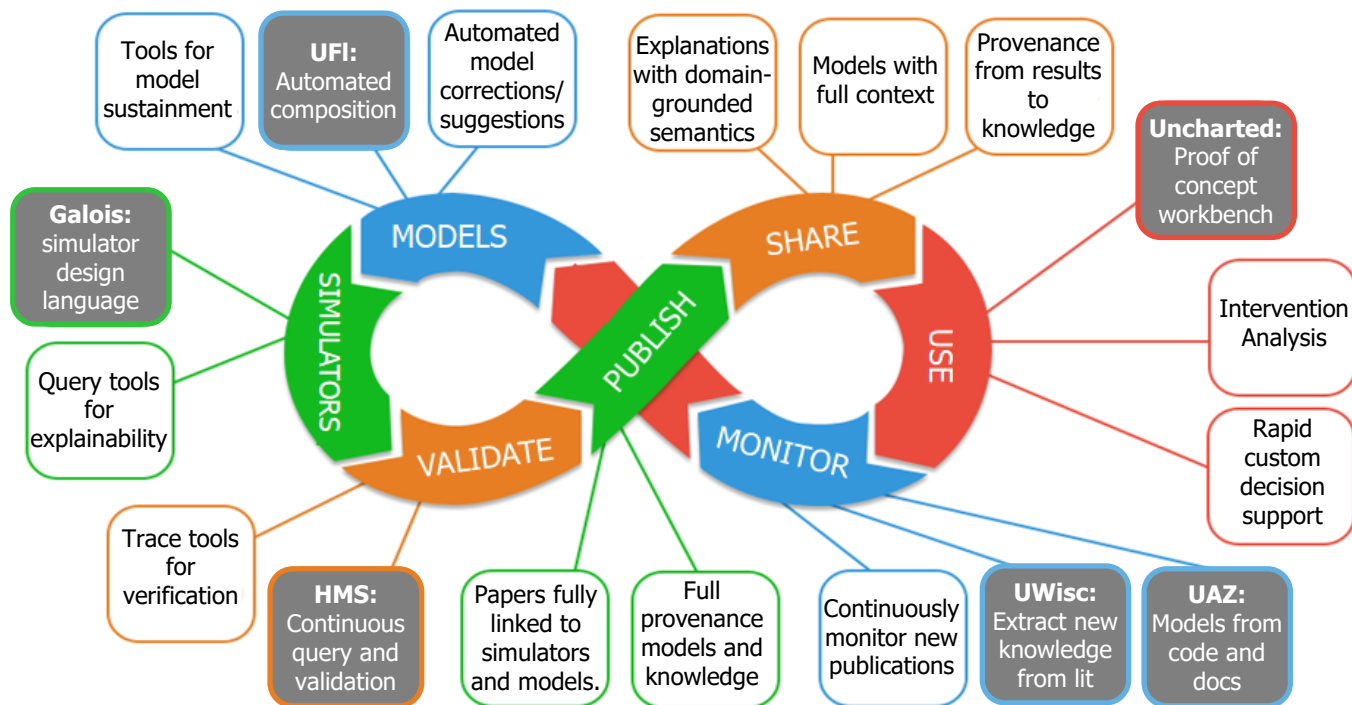
ASKEM for data and models

Enable continuous, traceable, and verifiable improvement of scientific modeling and publications



DevOps tools for software dev	Attributes	ASKEM integrating principles
from platform, network, OS, API, etc.	Abstract	from code, math framework, language, platform, etc.
distinct functional components	Compose	distinct model and simulator components
all suitable operations so developers can focus on higher-level functions	Automate	all suitable operations so scientists can focus on higher-level analysis
many complex components robustly	Orchestrate	complex components with diverse semantics (knowledge through execution)

ASKE AIE demonstrated portions of the continuous cycle



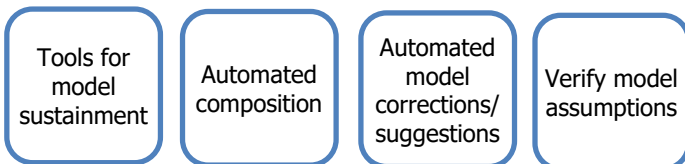
■ Demonstrated in ASKE AIE

Accomplishments, demonstrated in COVID-19 use-cases including molecular biology and epidemiology

- **Automatically extracted models from code**
 - Identified terms in literature but missing in code
 - Discovered faulty model assumptions
 - Extracted higher-level structure such as I/O, models, and solvers
 - Parsed equations from documents into comparable canonical representations
 - Enabled alignment of code with documentation and verification of consistency
- **Demonstrated machine assisted model modification**
 - Developed multi-model comparison
 - Composed models (e.g. epidemic with spatial and demographic)
 - Achieved 8x-48x speedup on computational fluid dynamics and epidemic modeling
- **Demonstrated automated repeatable skill assessment**
 - Demonstrated execution of historic model versions against new data
 - Developed reusable simulator meta-graph in abstract design language

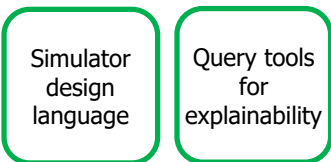
TA2: Machine-assisted modeling

Compose, extend, automate, and sustain models

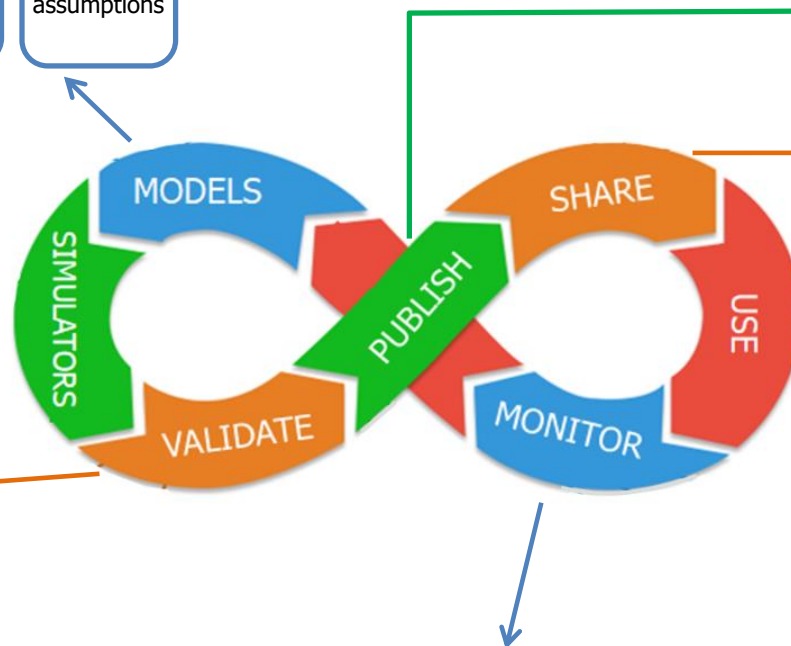


TA3: Machine-assisted simulators

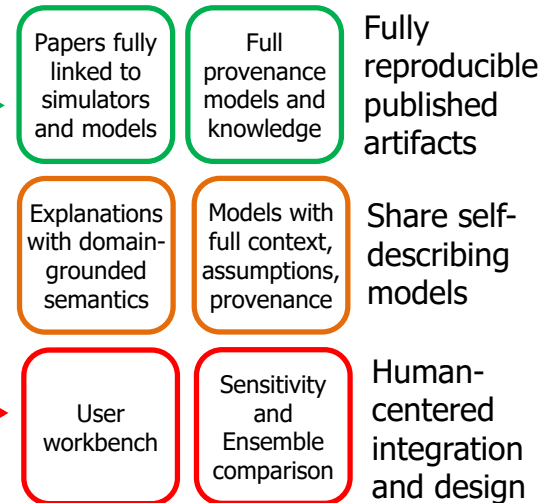
Abstract, compose, automate and optimize simulators



Automate continuous validation and fitness



TA4: Workbench for HMI and Integration



TA1: Machine-assisted knowledge discovery and curation

Track literature for updated knowledge



Attributes	Integrating principles
Abstract	Build abstractions for code, models, math frameworks, languages, and platforms
Compose	Compose distinct model and simulator components
Automate	Automate operations so scientists can focus on higher-level analysis
Orchestrate	Integrate components w/ diverse semantics (knowledge through execution)

Synthesize knowledge from available sources and modes for model recovery and annotation

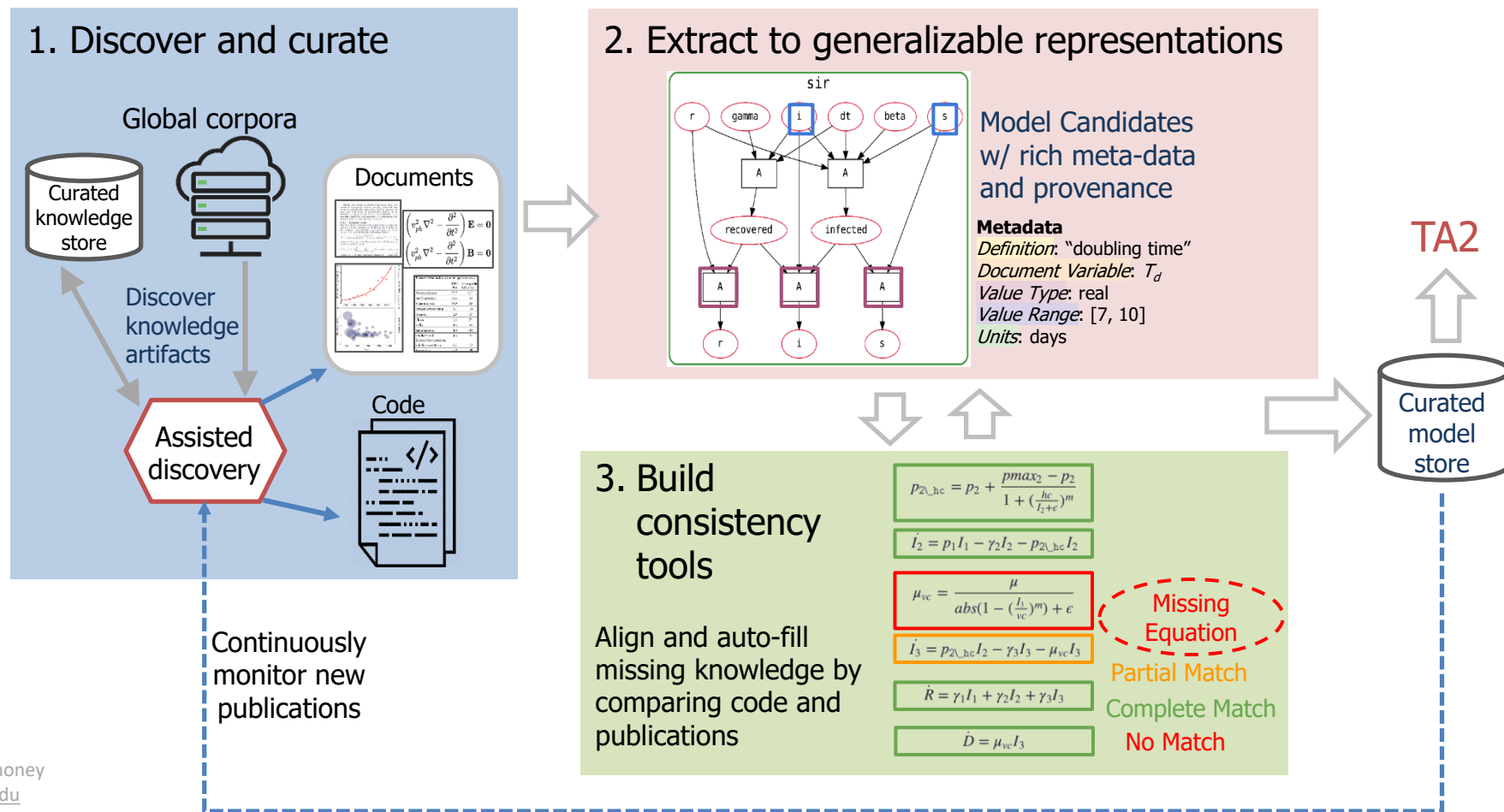
Challenges

1. Auto-capture provenance and context in arbitrary domains
2. Automated code analysis in diverse program languages to extract models and solvers with rich meta-data
3. Continuous alignment of models and (extracted and pushed) knowledge over the life-cycle

Potential Approaches

1. Extraction from multi-modal objects
 - Neuro-symbolic methods
2. Methods for static and dynamic code analysis
 - Identify models from code and documents
3. Partial graph matching
 - Equation alignment

1. Marius: Learning Massive Graph Embeddings on a Single Machine. Mahoney et al. (2021) <https://arxiv.org/abs/2101.08358> also <https://cosmos.wisc.edu>
2. AutoMATES: Automated Model Assembly from Text, Equations, and Software. A. Pyarelal et al. (2019). MOMACS. <https://arxiv.org/abs/2001.07295>
3. MathAlign: Linking Formula Identifiers to their Contextual Natural Language Descriptions. Alexeeva et al (2020). Proc.12th Lang. Res. Evaluation <https://aclanthology.org/2020.lrec-1.269/#>



Registered queries to provide updates for new knowledge artifacts related to models or simulators

Enable Framework-agnostic model assembly, comparison, and sustainment

Challenges

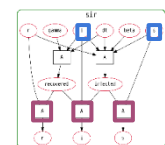
1. Automatically compose/decompose model components to accelerate update, customization, and sustainment
2. Learn to fill missing structure in model candidates. Auto-correction of models
3. Automated comparison of N models for discovery improvement, and causal ensemble explanations.

Potential Approaches

1. Causal Modularity and Algebraic Composition
2. Structure learning
3. Model-similarity assessment
 - E.g., graph matching and assumption variations

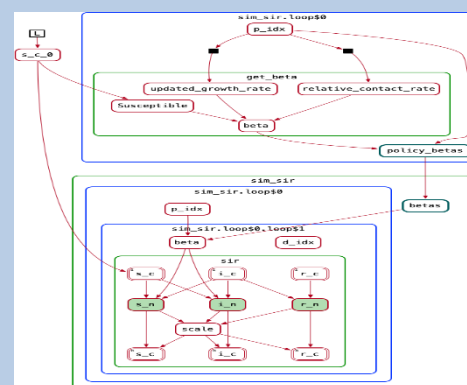
1. Compositional Scientific Computing with Catlab and SemanticModels. Fairbanks et al (2020). <https://arxiv.org/abs/2005.04831>
2. TASO: Optimizing Deep Learning Computation with Automatic Generation of Graph Substitutions. Jia et al. (2019) <https://dl.acm.org/doi/pdf/10.1145/3341301.3359630>
3. E. Davis, et al, "Machine-Assisted Extraction of Formal Semantics from Domain Specific Semi-Formal Diagrams," In the Proceedings of Modeling the World's Systems Conference. 2019

TA1 output: Model candidates/components

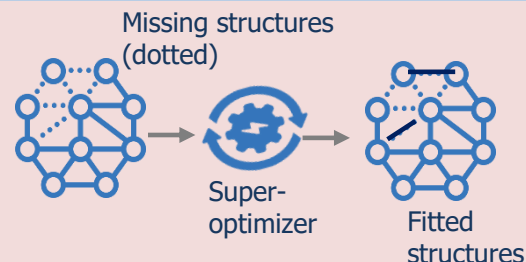


Select model(s)

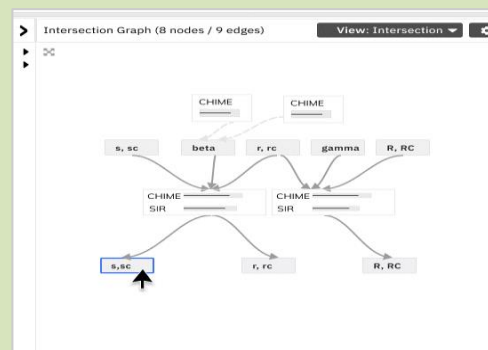
1. Automatically compose model components



2. Model corrections, learn to fill missing structure



3. Automated comparison of model candidates

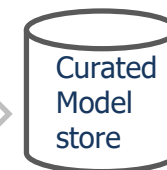


Register queries, discover new knowledge, data, etc.



TA1 Feedback loop

Sustainment



TA3

Automate simulator generation for complex, multi-model workflows and validation

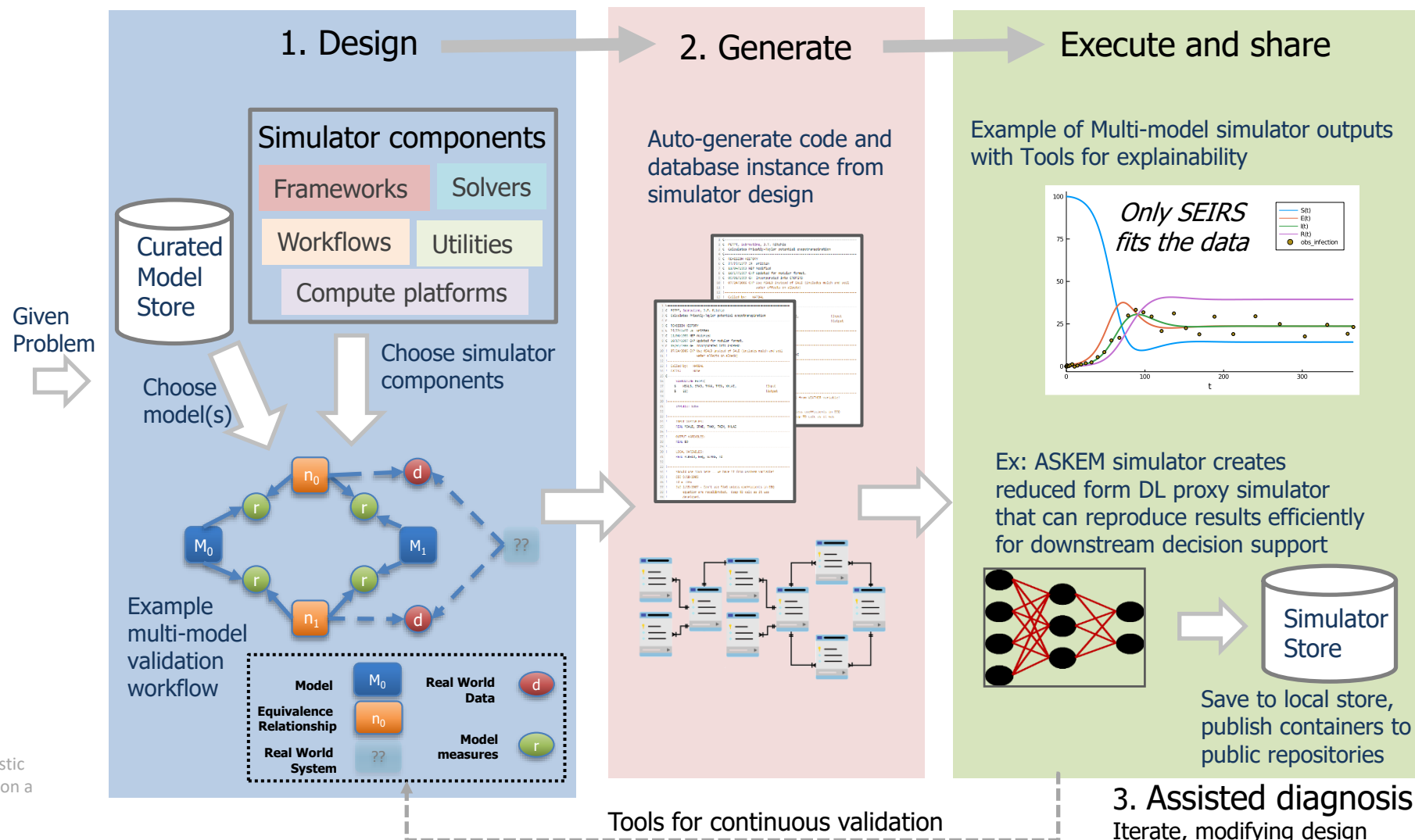
Challenges

1. Machine-assisted simulator design tools with model and knowledge provenance for explainability
2. Automated code generation for agile and adaptable simulators
3. Automated diagnostics of simulator fitness for continuous validation

Potential Approaches

1. Recommender system for component and architecture selection
2. Methods to optimally distinguish competing simulations
3. Iterated experimentation with causal identification and re-selection

1. AI in recommender systems. Zhang et al (2020). <https://link.springer.com/article/10.1007/s40747-020-00212-w>
2. Gutierrez, J. Extension of bisimulation, coinduction, and probabilistic bisimulation to allow for model-checked and black-box comparison on a measure bisimilarity basis. (2013): 108-110. https://www.mcr12.org/web/user_manual/index.html
3. Thompson, Michael, and Tom Fearn. "What exactly is fitness for purpose in analytical measurement?" Analyst 121.3 (1996): 275-278.



Build a visual meta-modeling platform for integrated modeling and simulation

Challenges

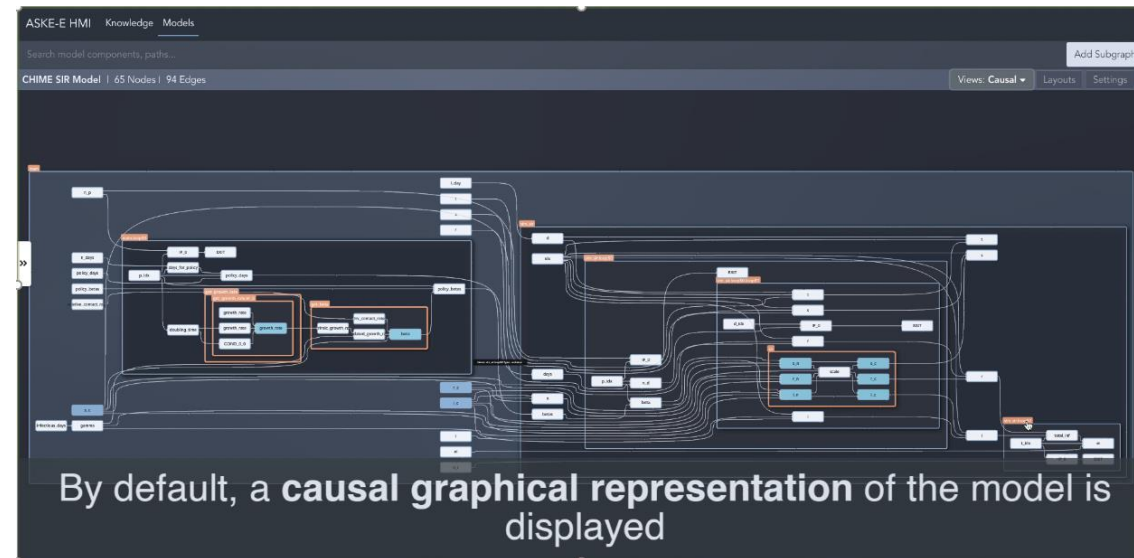
1. Model abstraction visualization
 - Spectrum of computational models and few common abstractions
2. Model comparison
 - Current workflows typically involve a single model
3. Knowledge exploration, traceability
4. Scalability in model exploration
 - Number of models and size/complexity of models
 - Visualization approaches

Potential Approaches

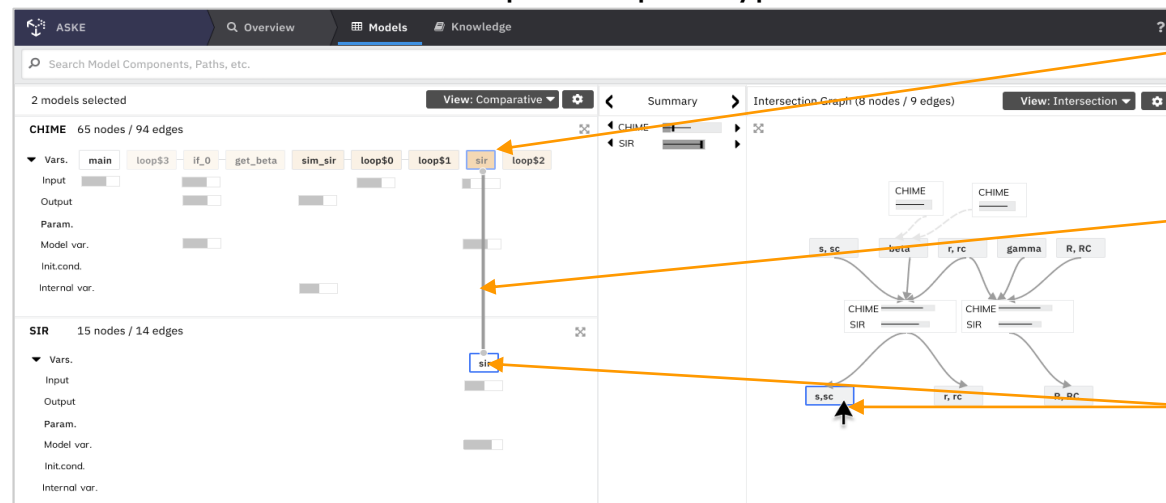
1. Link knowledge artifacts to model components
 - Assumptions, experiments, literature
2. Scalable comparison
 - Model structures, parameter surfaces, analysis of outputs
3. Complexity management
 - Attentional cues and system guidance
 - Identify related documents, parameters, and subgraphs

1. A multi-scale Approach for Biological Graph Visualization: Local Analysis in Global Context. Husain, F et al. (2021). Bio. Data Vis. (ISCB 2021). <https://arxiv.org/abs/2109.06828>

Model, meta-data, and linked knowledge exploration prototypes in ASKE AIE



Visualization of model comparison prototype in ASKE-E



Linear layout displays plates for each model

Plates are explicitly linked if they have overlapping components (nodes and edges)

Clicking shared variable in the intersection graph highlights the corresponding plates



Program Metrics

The ASKEM measurement approach leverages the lessons learned from ASKE-AIE to create metrics to drive all 5 program goals, tracked through continuous assessments by the T&E team and exercises with partners conducted at the end of each phase.

	TA1	TA2	TA3
Goals	Metrics Targets (AIE, Phase 1, Phase 2)	Metrics Targets (AIE, Phase 1, Phase 2)	Metrics Targets (AIE, Phase 1, Phase 2)
Accuracy		a1: Accuracy of a single model * (see note below)	a2: Multi-model ensemble accuracy *
Timeliness	t1: Time for accurate model extraction (from code/documents) (nm, 10x, 50x)	t2: Time to extend/modify model (10x, 50x, >200x)	t3: Time to create simulators for validation and prediction (nm, 50x, >200x)
Maintainability	t1 for hand-off to new owner (nm, 10x, 50x)	t2 for hand-off to new owner (nm, 10x, 50x)	t3 for hand-off to new owner (nm, 10x, 50x)
Generalizability [#]	Semantic domains supported (2, 4, 8)	Mathematical frameworks supported (2, 4, 6)	Simulator components supported (3, 10, 100)
Scalability [#]	Extracted model size (Number of variables) (10s, 100s, 1000s [†])	Number of components that can be (de)composed (2, 5, 10)	Simulator configurations automatically explored (nm, 100, 1000)

Notes and definitions:

nm = not measured

* At hackathon 1 the T&E team will define and baseline measures of forecasting skill (for predictions from models and ensembles) and fitness-for-purpose (for conditionals and counterfactuals) for COVID-19 models. Analogous measures for the Space Weather domain will be created at hackathon 2. The PAD will then be updated with targets for Phase 1 and 2.

10x = Activities that take "days to weeks" with current tools improved to "hours to days"

50x = Activities that take "weeks to months" with current tools improved to "hours to days"

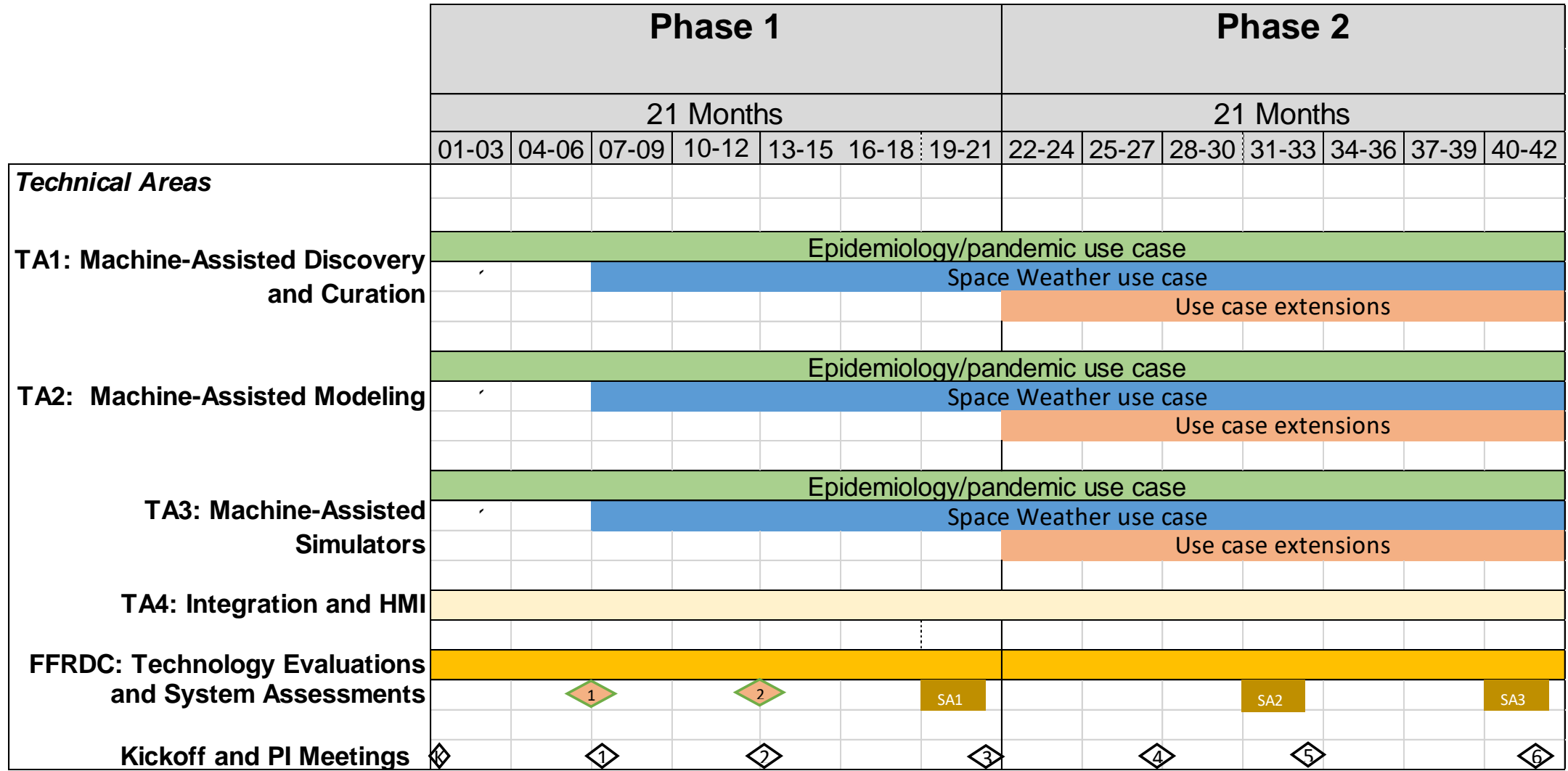
>200x = Activities that take years or are too expensive or impractical to do at all can be done in "days to weeks"

† As a simple example, a "very complicated" SEIR model with populations broken out by 8 demographic groups and 50 states could have 1600 variables

The Phase 1 targets for generalizability and scalability are estimated based on what we believe will be minimally necessary to succeed on the COVID-19 and SW use-cases.



Program Schedule



Technology Evaluation



Integrated System Evaluation



Proposal Information

- Performers may submit multiple proposals.
- Each proposal may address a single TA, combined TA1/TA2, or combined TA2/TA3. Other combinations of TAs will not be accepted.
- Proposers selected by the government for TA4 cannot be selected as prime or subcontractors on any other proposals.



Program deliverables and other responsibilities

- Executed Associate Contractor Agreement prior to kickoff
 - Program Kickoff Brief
 - Monthly financial status reports
 - Quarterly technical progress reports
 - Final Report
-
- Kickoff and PI Meeting attendance
 - Hackathon attendance and participation



QUESTIONS

- Please submit all questions via email to ASKEM@darpa.mil
- We will return to this Zoomgov at 1:15 pm EST with responses



Abstracts

- A. Cover Sheet (required): Include the administrative and technical points of contact (title, name, address, phone, e-mail, lead organization). Also include the BAA number, title of the proposed project (not the BAA title), Technical Area, subcontractors, estimated cost, duration of the project, and the label "ABSTRACT."
- B. Executive Summary: Clearly describe what is being proposed and what difference it will make (qualitatively and quantitatively).
- D. Technical Plan: Outline the technical challenges inherent in the approach and possible solutions for overcoming potential problems.
- E. Management and Capabilities Plan: Identify the principal investigator and provide a brief summary of expertise of the team.
- F. Cost and Schedule: Provide a rough cost estimate for resources over the proposed timeline of the project, broken down by phase and major cost items (e.g., labor, materials, etc.). Include cost estimates for each potential subcontractor (it may be a rough order of magnitude).



ASKEM - Agenda

Proposers Day – December 8, 2021

TIME (ET)	EVENT	SPEAKER
11:45 AM -	Submit Questions to ASKEM@darpa.mil (Webinar Chat and Q&A have been disabled for participants)	
11:45 AM - 1:15 PM	Informal Teaming Discussions while government reviews questions Slack registration at https://www.schafertmd.com/DARPA/I2O/ASKEM/PD/?p=teaming	
1:15 PM - 2:00 PM	Government responses to questions	(Answer attendee questions)



www.darpa.mil